Symbolic Regression Is All You Need: From Simulations to Scaling Laws in Binary Neutron Star Mergers

Phelipe Darc*
Artificial Intelligence for Physics Laboratory (Lab-IA)
and Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brazil.

Clecio R. Bom
Artificial Intelligence for Physics Laboratory (Lab-IA)
and Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brazil.

Charles Kilpatrick
Center for Interdisciplinary Exploration and Research in Astro- physics (CIERA)
and Department of Physics and Astronomy
North-western University, Evanston, IL 60208, USA.

Bernardo M. O. Fraga
Artificial Intelligence for Physics Laboratory (Lab-IA)
and Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brazil.

Gabriel S. M. Teixeira
Artificial Intelligence for Physics Laboratory (Lab-IA)
and Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brazil.

Abstract

Gravitational wave sources with electromagnetic counterparts have highlighted the need for predictive, interpretable models linking the parameters of compact binary systems to post-merger remnants and mass outflows. In this work, we explore AI-driven symbolic regression (SR) frameworks to derive updated analytical relations for disk ejecta mass in binary neutron star mergers, trained on state-of-the-art numerical relativity simulations. Our method reveals a set of compact equations that outperform existing fitting formulae across multiple statistical metrics while remaining physically interpretable. Notably, SR also enables alternative predictor sets (e.g., $\{M_1, M_2, \tilde{\Lambda}\}$) that match or exceed the accuracy of models relying solely on compactness of the lightest neutron star (C_1), enabling new parameter constraints from electromagnetic observations. Unlike traditional black-box machine learning models, these closed-form expressions generalize robustly to regions of the parameter space not represented in the training data, offering a physics-informed tool for multimessenger observations and constraints on the neutron star equation of state.

Machine Learning and the Physical Sciences Workshop, NeurIPS 2025.

1 Introduction

The coalescence of binary neutron stars (BNS) and black hole–neutron star (BHNS) systems are prime sources of gravitational waves (GWs) observed by the Advanced LIGO and Virgo detectors [1, 2]. These events are often accompanied by electromagnetic (EM) counterparts, including kilonovae. Kilonovae are powered by the radioactive decay of *r*-process elements synthesized in neutron-rich ejecta, whose brightness and evolution depend on their total dynamical ejecta mass ($M_{\rm dyn}$), the mass remaining bound in the remnant accretion disk ($M_{\rm disk}$), and outflows from the disk.

Kilonova observations provide a unique window into the physical conditions of the merger remnant, including constraints on the mass, composition, and geometry of the ejected material. Accurately modeling the ejecta properties is essential for applying EM signals to infer key source parameters, including the component neutron star (NS) masses (M_1, M_2), compactness (C_1, C_2), and ultimately the equation of state (EoS) of dense nuclear matter. The nuclear EoS describes the properties of dense matter, which is fundamental for understanding the behavior of NS throughout astrophysics.

The connection between the EoS parameters and ejecta properties is informed by numerical relativity (NR) Simulations. However, generating such simulations is computationally expensive and highly sensitive to the EoS adopted. Over the years, numerous studies have proposed distinct scaling relations to model key properties of BNS mergers, such as the accretion disk mass ($M_{\rm disk}$), the dynamical ejecta mass ($M_{\rm dyn}$), and the average ejecta velocity, as functions of underlying EOS parameters calibrated on existing NR simulation datasets.

Radice et al. (2018) [21] first identified a correlation between the mass of the remnant accretion disk and the binary tidal deformability. Coughlin et al. (2019) [7] incorporated the threshold mass $M_{\rm thr}$, above which the remnant promptly collapses to a BH, while Dietrich et al. (2020) [9] extended this approach with a larger dataset, introducing explicit dependence on the mass ratio q. Krüger and Foucart (2020) [14] highlighted the dominant role of the lighter star's compactness C_1. Building on broader NR compilations, Nedora et al. (2021) [18] proposed polynomial fits in q and $\tilde{\Lambda}$. More recently, Lund et al. (2025) [15] reaffirmed the strong C_1 dependence with an updated formula calibrated on the most extensive dataset to date [5, 6, 20, 18, 14].

Despite their widespread use, analytical fits for neutron star merger ejecta face serious limitations in reliability and generalization. Differences between models often exceed reported uncertainties, extrapolation beyond calibration can produce unphysical results, and uncertainty quantification remains limited [10]. These fits are also highly sensitive to the adopted microphysics and EOS [18], and commonly used dynamical ejecta models fail for GW190425-like systems [6], underscoring the need for more robust approaches. In this work, we present symbolic regression (SR) algorithms [16] that can autonomously discover analytical expressions that provide interpretability and better generalization across data sets.

2 Symbolic Regression & Numerical Relativity Simulation Dataset

We focus our analysis specifically on the $M_{\rm disk}$, which represents a particularly challenging quantity to estimate due to its reported values from simulations [10] being highly sensitive to the time at which the measurement is performed, as well as to the criteria used to distinguish the accretion disk from the central remnant in neutron star–disk systems.

Data: We adopt two primary datasets of NR simulations. For training (calibration), we employ the same dataset utilized in [14] (hereafter Krüger20), which includes 56 NR simulations from [21] and [12]. This dataset spans disk masses in the range 10^{-4}–$0.234\,M_\odot$ and mass ratios $q = 0.77$–1.0. For evaluation, we use the NR simulation compilation from [18] (hereafter Nedora21), which includes all simulations from the [14] dataset along with 63 additional simulations from multiple sources. This expanded test set covers a wider physical regime, with disk masses ranging from 3×10^{-5}–$0.30\,M_\odot$ and mass ratios from 0.55–1.0. The analytical fitting formulae proposed by `Radice18` [21], `Coughlin19` [7], `Dietrich20` [9], `Krüger20` [14], and `Lund25` [15] are provided in Appendix 4 [1]. The `Nedora21` dataset provides only $M_{1,2}$, $C_{1,2}$, and Λ; therefore, we were unable to apply this test

[1] Throughout this paper, we label the two neutron stars with subscripts 1 and 2. The individual gravitational masses are denoted by M_1 and M_2, while the corresponding baryonic masses are written as M_{b1} and M_{b2}. The total gravitational mass is defined as $M_{\rm tot} = M_1 + M_2$, and the mass ratio is given by $q = M_1/M_2 < 1.0$.

set to some of the fitting formulas proposed in the literature. This limitation will be addressed in future work.

Methods: Symbolic regression is a subfield of machine learning and genetic programming (GP) that aims to infer interpretable, closed-form mathematical expressions directly from data [3, 13, 22]. Unlike traditional regression techniques, which optimize parameters within a fixed model structure, SR searches over both the space of model structures and parameter values. Learning models in the form of simple mathematical expressions offers much more than just potentially improved predictive power [17] but it also enhances human interpretability. Most SR algorithms start with a population of random expressions iteratively refined through mutation, crossover, simplification, and constant optimization [8]. Models are typically evaluated with mean squared error under multi-objective optimization, balancing accuracy and simplicity until convergence, producing a set of candidate expressions ranked by error and complexity, where each operator, variable, or constant contributes one unit to the total complexity by default.

SR Training Setup: We evaluated two symbolic regression frameworks for disk mass prediction. `PyOperon` [4] is a Python wrapper to Operon, a popular SR method that uses genetic programming to explore a hypothesis space of possible symbolic expressions; we used the official implementation (link) with default hyperparameters and operators `{+, -, *, /, sin, sqrt, log, tanh}`. `PySR` [8] is an evolutionary SR framework that balances accuracy and complexity via a multi-objective loss, with GPU and multi-threading support; we used the official implementation (link). For `PySR`, we tested three operator sets: Core (`{+, -, *, /; sqrt, log, exp, abs}`), Extended (`{+, -, *, /, max, min; sqrt, log, exp, square, cube, abs}`), and All (`{+, -, *, /, max, min, pow; sqrt, log, exp, square, cube, abs, sin, cos, tan, sinh, cosh, tanh}`).

We also explored a physics-informed approach using predefined templates with the post-merger disk mass written as a linear combination of symbolic sub-expressions tied to variable groups:

$$M_{\mathrm{disk,pred}} = a \cdot f(C_1, M_1) + b \cdot g(C_2, M_2) + c \cdot h(q, \tilde{\Lambda}) + d, \tag{1}$$

where a, b, c, d are free coefficients and f, g, h are symbolic functions discovered by SR. This expression allows us to isolate and quantify the influence of physical parameters on the remnant disk mass. We additionally queried `ChatGPT-4.0` (see Appendix 4) for an alternative template inspired by literature fits:

$$M_{\mathrm{disk,pred}} = a_1 f(q_0, C_1) + a_2 g(C_1, \tilde{\Lambda}) + a_3 h(M_1, M_2) + a_4 t(q_0) + a_5, \tag{2}$$

where $a_{1\ldots5}$ are scalar coefficients and f, g, h, t are symbolic functions learned by SR. Their structure draws inspiration from previous fitting relations such as `Krüger20`, `Radice18`, and `Nedora21`. All models were calibrated on Krüger20 dataset for 3000 iterations with a maximum expression size of 20 (10 for the physics-informed template) and tree depth up to 5.

3 Results

Model Selection: For the `PySR` approach, models were selected along the Pareto front, which reflects the trade-off between predictive accuracy and complexity. From each training configuration, we chose a single representative model with complexity below 10, prioritizing those whose Mean Squared Error (MSE) remained stable as complexity increased, indicating robustness to overfitting. The resulting equations are listed in Appendix 4, each labeled with an identifier encoding its configuration (e.g., `PySR_template_ext` refers to a model trained on Krüger20). In the template-based approach, we restricted the selection to models with complexity ≤ 20, since the physics-informed template itself carries a baseline complexity of 13 (17 for the LLM-based template).

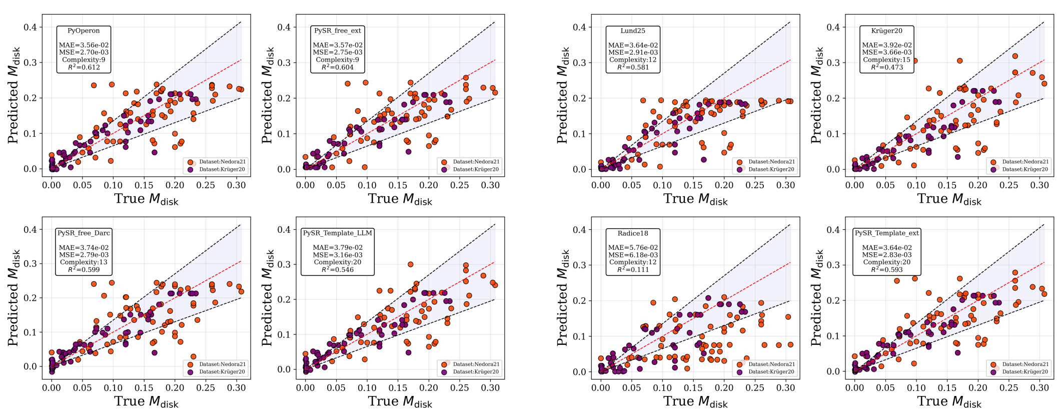
For the `PyOperon` framework, we selected the expression with a complexity of 9, chosen for its lower structural complexity and exclusive dependence on the compactness of the lighter neutron star (C_1):

The dimensionless tidal deformability is denoted by $\tilde{\Lambda}$, and the compactnesses of the two stars are represented by C_1 and C_2.

Table 1: Quantitative performance comparison of symbolic regression models and literature fitting formulae for predicting post-merger disk mass. Metrics include MSE, MAE, coefficient of determination (R^2), and Bayesian Information Criterion (BIC), along with the model complexity (Comp.) and dependence on specific physical parameters. All SR models were calibrated on the `Krüger20` dataset, and all fitting formulae were evaluated on the `Nedora21` test set.

Model	MSE	MAE	R^2	BIC	Comp.	Dependence
PyOperon	0.002697	0.035435	0.611964	-660.945373	9	C_1
PySR_Free_Ext	0.002751	0.035714	0.604268	-658.608259	9	C_1, M_1
PySR_Free_M12	0.002789	0.037436	0.598510	-637.772776	13	M_1, M_2, $\tilde{\Lambda}$
PySR_Template_Ext	0.002827	0.036365	0.593230	-602.764208	20	C_1, C_2, $\tilde{\Lambda}$
Lund25	0.002910	0.036369	0.581383	-637.580875	12	C_1
PySR_Free_All	0.003029	0.036696	0.564164	-651.900268	8	C_1
PySR_Template_LLM	0.003156	0.037854	0.545986	-589.732988	20	C_1, M_2
PySR_Template_All	0.003322	0.037620	0.522114	-593.150872	20	C_1, $\tilde{\Lambda}$
PySR_Free_Core	0.003340	0.038944	0.520834	-640.621426	8	C_1
Krüger20	0.003665	0.039238	0.472704	-595.777066	15	C_1, M_1
PySR_Template_Core	0.005263	0.046414	0.242850	-528.877037	20	M_1, M_2, q, $\tilde{\Lambda}$
Radice18	0.006179	0.057595	0.111014	-524.063305	12	$\tilde{\Lambda}$

$$M_{disk}^{PyOperon} = 0.118824 - (0.142985 \times \sin(\sin(40.896317 \times C_1))) \tag{3}$$



(a) (b)

Figure 1: Predicted disk mass versus disk mass from numerical relativity simulations. The left panel (a) presents results obtained with symbolic regression expressions, while the right panel (b) shows results from literature fitting formulae together with the `PySR` Template expression proposed by the LLM. Purple points correspond to the calibrated regime (Krüger20 dataset), whereas orange points represent the uncalibrated regime (Nedora21 dataset). Each inset box reports the model name, mean absolute error, mean squared error, expression complexity, and coefficient of determination (R^2).

Model Evaluation: We compare literature fitting formulae with SR-derived expressions to evaluate whether SR models can (i) outperform existing fitting formulae in predictive accuracy, (ii) generalize beyond their calibration regime, (iii) recover physically meaningful dependencies, and (iv) produce more compact expressions while maintaining high accuracy. Model performance is assessed using mean absolute error (MAE), mean squared error (MSE), the Bayesian Information Criterion (BIC), and the coefficient of determination (R^2). Table 1 summarizes the performance metrics, complexity, and parameter dependencies for all SR models trained on the `Krüger20` dataset, together with three representative literature formulae: `Krüger20`, `Radice18`, and `Lund25`. Nearly all SR-derived scaling relations outperform the `Krüger20` and `Radice18` fits on the `Nedora21` dataset, confirming that literature models often fail to generalize beyond their calibration regime [10]. Notably, four of the nine SR expressions also surpass `Lund25` across all metrics; in particular, `PyOperon` achieves a BIC more than 23 points lower, providing strong statistical evidence for the SR approach, even though it is evaluated on unseen data while Lund25 is tested on its calibration set.

Although trained on the full set of available input variables $(M_{1,2}, C_{1,2}, \tilde{\Lambda})$, SR consistently rediscovers the dominant influence of the lighter star's compactness (C_1), and in some cases also its mass (M_1), demonstrating both interpretability and robustness. Finally, Figures 1a and 1b compare predicted and true $M_{\rm disk}$, with purple markers denoting data beyond the calibration regime for the SR expressions. While Lund25 is unable to reproduce cases with $M_{\rm disk} > 0.20\, M_\odot$, models such as PySR_Free_M12 and PySR_Template_LLM generalize effectively to this high-mass regime. In the intermediate disk-mass regime ($M_{\rm disk} \sim 0.05$–$0.20\, M_\odot$), the top SR models—PyOperon and PySR_Free_Ext—achieve MSE values of 2.12×10^{-3} and 2.09×10^{-3}, outperforming Krüger20 (3.31×10^{-3}) and matching Lund25 (1.95×10^{-3}). This underscores SR's ability to generalize beyond the training domain. Furthermore, SR can discover alternative predictor sets with comparable performance: e.g., PySR_Free_M12, which depends on $(M_1, M_2, \tilde{\Lambda})$, performs as well as Lund25, which relies solely on C_1. This flexibility enables the possibility of constraining previously inaccessible EOS parameters through disk mass measurements.

In this work, we focused our analysis on performance metrics and model complexity. However, understanding the physical meaning of these models in extreme regimes is equally important. Since the existing fitting formulas are widely used in multi-messenger analyses, identifying new and reliable scaling relations requires exploring their qualitative behavior for large neutron star radii and high compactness, where numerical simulations are still limited. This aspect will be further investigated in the final version of this work. We suggest using the PyOperon model as an alternative to existing fitting formulas in the literature, given its strong performance and its dependence solely on the compactness parameter C_1.

4 Conclusion

Symbolic regression offers a promising framework for bridging data-driven modeling with scientific understanding. In this ongoing work, we evaluate two SR frameworks,PySR and PyOperon, using both physics-informed templates and free-form discovery to model post-merger disk mass in BNS mergers. Trained on a small dataset (56 NR simulations from [14]), the resulting expressions generalize effectively to the broader [18] compilation, demonstrating SR as an accurate and interpretable tool for uncovering physically meaningful relations in high-dimensional astrophysical data. Notably, models such as PySR_Free_M12, which depend on M_1, M_2, and $\tilde{\Lambda}$, achieve accuracy comparable to the literature while potentially enabling new parameter constraints from electromagnetic observations.

In future work, we plan to extend this study by incorporating additional SR algorithms and methodologies, such as AI-Feynman[23] and SciMED [11]. We also aim to calibrate these expressions on a more diverse and carefully curated dataset, including a broader range of NR simulations with more detailed physics. Another important direction involves applying SR to model the dynamical ejecta mass, with the goal of improving constraints on the neutron star equation of state from events such as GW170817 [19].

References

[1] J Aasi et al. "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7 (Mar. 2015), p. 074001. ISSN: 1361-6382. DOI: 10.1088/0264-9381/32/7/074001. URL: http://dx.doi.org/10.1088/0264-9381/32/7/074001.

[2] F Acernese et al. "Advanced Virgo: a second-generation interferometric gravitational wave detector". In: *Classical and Quantum Gravity* 32.2 (Dec. 2014), p. 024001. ISSN: 1361-6382. DOI: 10.1088/0264-9381/32/2/024001. URL: http://dx.doi.org/10.1088/0264-9381/32/2/024001.

[3] Dimitrios Angelis, Filippos Sofos, and Theodoros E. Karakasidis. "Artificial Intelligence in Physical Sciences: Symbolic Regression Trends and Perspectives". In: *Archives of Computational Methods in Engineering* 30.6 (Apr. 2023), pp. 3845–3865. ISSN: 1886-1784. DOI: 10.1007/s11831-023-09922-z. URL: http://dx.doi.org/10.1007/s11831-023-09922-z.

[4] Bogdan Burlacu, Gabriel Kronberger, and Michael Kommenda. "Operon C++: an efficient genetic programming framework for symbolic regression". In: *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*. GECCO '20. ACM, July 2020, pp. 1562–1570. DOI: 10.1145/3377929.3398099. URL: http://dx.doi.org/10.1145/3377929.3398099.

[5] Alessandro Camilletti et al. "Geometric and thermodynamic characterization of binary neutron star accretion discs". In: *Physical Review D* 109.6 (Mar. 2024). ISSN: 2470-0029. DOI: 10.1103/physrevd.109.063023. URL: http://dx.doi.org/10.1103/PhysRevD.109.063023.

[6] Alessandro Camilletti et al. "Numerical relativity simulations of the neutron star merger GW190425: microphysics and mass ratio effects". In: *Monthly Notices of the Royal Astronomical Society* 516.4 (Aug. 2022), pp. 4760–4781. ISSN: 1365-2966. DOI: 10.1093/mnras/stac2333. URL: http://dx.doi.org/10.1093/mnras/stac2333.

[7] Michael W Coughlin et al. "Multimessenger Bayesian parameter inference of a binary neutron star merger". In: *Monthly Notices of the Royal Astronomical Society: Letters* 489.1 (Aug. 2019), pp. L91–L96. ISSN: 1745-3933. DOI: 10.1093/mnrasl/slz133. URL: http://dx.doi.org/10.1093/mnrasl/slz133.

[8] Miles Cranmer. *Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl*. 2023. DOI: 10.48550/ARXIV.2305.01582. URL: https://arxiv.org/abs/2305.01582.

[9] Tim Dietrich et al. "Multimessenger constraints on the neutron-star equation of state and the Hubble constant". In: *Science* 370.6523 (Dec. 2020), pp. 1450–1453. ISSN: 1095-9203. DOI: 10.1126/science.abb4317. URL: http://dx.doi.org/10.1126/science.abb4317.

[10] Amelia Henkel et al. "Study of the agreement between binary neutron star ejecta models derived from numerical relativity simulations". In: *Physical Review D* 107.6 (Mar. 2023). ISSN: 2470-0029. DOI: 10.1103/physrevd.107.063028. URL: http://dx.doi.org/10.1103/PhysRevD.107.063028.

[11] Liron Simon Keren, Alex Liberzon, and Teddy Lazebnik. "A computational framework for physics-informed symbolic regression with straightforward integration of domain knowledge". In: *Scientific Reports* 13.1 (Jan. 2023). ISSN: 2045-2322. DOI: 10.1038/s41598-023-28328-2. URL: http://dx.doi.org/10.1038/s41598-023-28328-2.

[12] Kenta Kiuchi et al. "Revisiting the Lower Bound on Tidal Deformability Derived by AT 2017gfo". In: *The Astrophysical Journal Letters* 876.2 (May 2019), p. L31. ISSN: 2041-8213. DOI: 10.3847/2041-8213/ab1e45. URL: http://dx.doi.org/10.3847/2041-8213/ab1e45.

[13] JohnR. Koza. "Genetic programming as a means for programming computers by natural selection". In: *Statistics and Computing* 4.2 (June 1994). ISSN: 1573-1375. DOI: 10.1007/bf00175355. URL: http://dx.doi.org/10.1007/BF00175355.

[14] Christian J. Krüger and Francois Foucart. "Estimates for disk and ejecta masses produced in compact binary mergers". In: *Physical Review D* 101.10 (May 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.101.103002. URL: http://dx.doi.org/10.1103/PhysRevD.101.103002.

[15] Kelsey A. Lund et al. "Kilonova Emissions from Neutron Star Merger Remnants: Implications for the Nuclear Equation of State". In: *The Astrophysical Journal* 987.1 (June 2025), p. 56. ISSN: 1538-4357. DOI: 10.3847/1538-4357/add148. URL: http://dx.doi.org/10.3847/1538-4357/add148.

[16] Nour Makke and Sanjay Chawla. "Interpretable scientific discovery with symbolic regression: a review". In: *Artificial Intelligence Review* 57.1 (Jan. 2024). ISSN: 1573-7462. DOI: 10.1007/s10462-023-10622-0. URL: http://dx.doi.org/10.1007/s10462-023-10622-0.

[17] Madhav R. Muthyala et al. "SyMANTIC: An Efficient Symbolic Regression Method for Interpretable and Parsimonious Model Discovery in Science and Beyond". In: *Industrial amp; Engineering Chemistry Research* 64.6 (Feb. 2025), pp. 3354–3369. ISSN: 1520-5045. DOI: 10.1021/acs.iecr.4c03503. URL: http://dx.doi.org/10.1021/acs.iecr.4c03503.

[18] Vsevolod Nedora et al. "Mapping dynamical ejecta and disk masses from numerical relativity simulations of neutron star mergers". In: *Classical and Quantum Gravity* 39.1 (Dec. 2021), p. 015008. ISSN: 1361-6382. DOI: 10.1088/1361-6382/ac35a8. URL: http://dx.doi.org/10.1088/1361-6382/ac35a8.

[19] Vsevolod Nedora et al. "Numerical Relativity Simulations of the Neutron Star Merger GW170817: Long-term Remnant Evolutions, Winds, Remnant Disks, and Nucleosynthesis". In: *The Astrophysical Journal* 906.2 (Jan. 2021), p. 98. ISSN: 1538-4357. DOI: 10.3847/1538-4357/abc9be. URL: http://dx.doi.org/10.3847/1538-4357/abc9be.

[20] Vsevolod Nedora et al. "Spiral-wave Wind for the Blue Kilonova". In: *The Astrophysical Journal Letters* 886.2 (Nov. 2019), p. L30. ISSN: 2041-8213. DOI: 10.3847/2041-8213/ab5794. URL: http://dx.doi.org/10.3847/2041-8213/ab5794.

[21] David Radice et al. "Binary Neutron Star Mergers: Mass Ejection, Electromagnetic Counterparts, and Nucleosynthesis". In: *The Astrophysical Journal* 869.2 (Dec. 2018), p. 130. ISSN: 1538-4357. DOI: 10.3847/1538-4357/aaf054. URL: http://dx.doi.org/10.3847/1538-4357/aaf054.

[22] Michael Schmidt and Hod Lipson. "Distilling Free-Form Natural Laws from Experimental Data". In: *Science* 324.5923 (Apr. 2009), pp. 81–85. ISSN: 1095-9203. DOI: 10.1126/science.1165893. URL: http://dx.doi.org/10.1126/science.1165893.

[23] Silviu-Marian Udrescu and Max Tegmark. "AI Feynman: A physics-inspired method for symbolic regression". In: *Science Advances* 6.16 (Apr. 2020). ISSN: 2375-2548. DOI: 10.1126/sciadv.aay2631. URL: http://dx.doi.org/10.1126/sciadv.aay2631.

Appendix A: Fitting Formulae from the Literature

We summarize below the main empirical models proposed in the literature to estimate the post-merger accretion disk mass $M_{\rm disk}$ as a function of binary and neutron star properties.

Radice et al. (2018) [21]:

$$M_{\rm disk} = \max\left(10^{-3},\, \alpha + \beta \tanh\left(\frac{\tilde{\Lambda} - \gamma}{\delta}\right)\right), \tag{4}$$

where $\alpha = 0.084$, $\beta = 0.127$, $\gamma = 567.1$, and $\delta = 405.14$.

Coughlin et al. (2019) [7]:

$$\log_{10} M_{\rm disk} = \max\left(-3,\, a\left(1 + b\tanh\left(\frac{c - M_{\rm tot}/M_{\rm thr}}{d}\right)\right)\right), \tag{5}$$

where $a = -31.335$, $b = -0.9760$, $c = 1.0474$, and $d = 0.05957$.

Krüger & Foucart (2020) [14]:

$$M_{\rm disk} = M_1 \cdot \max\left(5 \times 10^{-4}, (aC_1 + c)^d\right), \tag{6}$$

with $a = -8.1324$, $c = 1.4820$, and $d = 1.7784$.

Dietrich et al. (2020) [9]:

$$q_{\rm si} = \frac{1}{2}\tanh\left[\beta(q_0 - Q_{\rm trans})\right], \tag{7}$$

$$a = a_0 + \Delta a \cdot q_{\rm si}, \qquad b = b_0 + \Delta b \cdot q_{\rm si}, \tag{8}$$

$$\log_{10} M_{\rm disk} = \max\left(-3,\, a\left(1 + b\tanh\left(\frac{c - M_{\rm tot}/M_{\rm thr}}{d}\right)\right)\right), \tag{9}$$

where the coefficients are $a_0 = -1.581$, $\Delta a = -2.439$, $b_0 = -0.538$, $\Delta b = -0.406$, $c = 0.953$, $d = 0.0417$, $\beta = 3.910$, and $Q_{\rm trans} = 0.9$.

Lund et al. (2025) [15]:

$$\log_{10} M_{\rm disk} = \alpha \tanh(\beta C_1 + \gamma) + \delta, \tag{10}$$

where $\alpha = -1.27$, $\beta = 68.01$, $\gamma = -11.72$, and $\delta = -1.98$.

Appendix B: Symbolic Regression-Derived Expressions

The following expressions were derived using symbolic regression calibrated on [14] dataset.

PySR_Free_Core (Complexity 8):

$$M_{\mathrm{disk}} = \left|\frac{0.1126}{C_1} - 0.6026\right| - 0.0177 \tag{11}$$

PySR_Free_Ext (Complexity 9):

$$M_{\mathrm{disk}} = \frac{\max\left[M_1(0.1750 - C_1),\, 0.00136\right]}{0.2308} \tag{12}$$

PySR_Free_All (Complexity 8):

$$M_{\mathrm{disk}} = \tanh\left(\max\left[0.00622, \frac{0.1763 - C_1}{0.1920}\right]\right) \tag{13}$$

PySR_Free_M12 (Complexity 13):

$$M_{\mathrm{disk}} = \min\left(0.2676, \frac{0.0006162 \cdot \tilde{\Lambda}}{M_1^3}\right) - \frac{0.1508}{M_2^3} \tag{14}$$

PySR_Template_LLM (Complexity 20, simplified to 16):

$$\begin{aligned} M_{\mathrm{disk}} &= 18170.047\, C_1^{6.911347} - 6.031138\, e^{C_1} \\ &- 0.02342195\, M_2 - 0.017214041 \times 432.73465 - 0.3142357. \end{aligned} \tag{15}$$

ChatGPT-4.0: Physics-Informed Symbolic Regression Prompt

We fed the [10] study to the ChatGPT memory and used the following prompt:

```
[Start of Task - reset all prior context]
You are a domain expert in relativistic astrophysics and gravitational
wave modeling. Use only information derived from published, peer-reviewed
physics literature related to binary neutron star (BNS) mergers and their
associated ejecta (e.g., kilonovae, accretion disk formation).

TASK: Based on the information from this article (and its cited references),
synthesize a new physically motivated Ansatz - that is, a symbolic regression
expression - for the accretion disk mass (M_{\mathrm{disk}}) formed after a BNS
merger. This Ansatz will be used in PySR as a symbolic regression template
function for fitting Numerical Relativity (NR) simulation outputs of BNS mergers.
Your goal is to infer M_{\mathrm{disk}} using a closed-form expression based on
physical parameters.

Requirements:
Input Variables: Use a subset or full set of these physically meaningful input
parameters:
    M_1: gravitational mass of neutron star 1 [solar masses]
    M_2: gravitational mass of neutron star 2 [solar masses]
    C_1: compactness of neutron star 1 [GM/(Rc^2)]
    C_2: compactness of neutron star 2
    q: mass ratio (M_2/M_1)
    \tilde{\Lambda}: reduced tidal deformability

Output Variable:
    The expression should predict:
```

```
    M_{\mathrm{disk}}: accretion disk mass [solar masses]

Return Format:
    Provide a complete PySR TemplateExpressionSpec object, e.g.,
    template = TemplateExpressionSpec(
        expressions=["f", "g", "h", "t"],
        variable_names=["M1", "M2", "C1", "C2", "q", "Lambda_tilde"],
        parameters={"p1": decide, "p2": decide},
        combine="template fitting function you created"
    )

Explanation Required:
    After providing the PySR template expression:
    - Justify the physical basis for the structure of each symbolic function
      (e.g., why q, \tilde{\Lambda}, and C are used together or multiplied).
    - cite examples or models from the literature
      (e.g., Radice+2018, Dietrich+2017, Krüger+2020, Kiuchi+2019, etc.)
      that support your formulation.
    - Discuss if the functional form is expected to extrapolate reasonably
      outside the training domain, or whether it's more phenomenological.

Your task is to derive a physics-informed symbolic expression for
M_{\mathrm{disk}}, grounded in literature, with a structure usable
in symbolic regression.
```

ChatGPT-4.0 produced the following answer:

```
Based on the detailed analysis and comparative study presented in Henkel et al. (2022),
and using insight from the fitting formulas in Krüger & Foucart (2020), Radice et al. (2018),
Dietrich et al. (2020), and Nedora et al. (2021), I propose the following physics-informed
Ansatz for the accretion disk mass M_{\mathrm{disk}} formed after a BNS merger.

Physical Motivation:
Accretion disk mass depends sensitively on:
    - The compactness C1 of the less massive neutron star - more compact stars
      resist disruption.
    - The mass ratio q = M2 / M1, with lower values (more unequal masses) leading
      to stronger tidal effects and larger disks.
    - The reduced tidal deformability \tilde{\Lambda}, which captures the binary's
      susceptibility to deformation.
    - A threshold behavior when approaching prompt collapse, typically captured
      via a tanh transition (e.g., Dietrich+2020).

Proposed PySR Template:
template = TemplateExpressionSpec(
    expressions=["f", "g", "h", "t"],
    variable_names=["M1", "M2", "C1", "C2", "q", "Lambda_tilde"],
    parameters={"p1": 3, "p2": 2},
    combine="p1[1]*f(q, C1) + p1[2]*g(Lambda_tilde)
             + p1[3]*h(M1, M2) + p2[1]*t(q) + p2[2]"
)
-----------------------------------------------------------------------------------------
```